Global Journal of Computer Science and Technology
Volume 11 Issue 11 Version 1.0 July 2011
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

# An N-Square Approach for Reduced Complexity Non-Binary Encoding

By G.Srinivas; P.V.G.D.Prasad Reddy; K.Ramya

*Abstracts -* There is always a need for the compression of data to facilitate its easy transmission and storage. Several lossy and lossless techniques have been developed in the past few decades. Lossless techniques allow compression without any loss of information. In this paper, we propose a new algorithm for lossless compression. Our experimental results show that the proposed algorithm performs compression in lesser iterations than the existing Non-Binary Huffman coding without affecting the average number of digits required to represent the symbols, thereby reducing the complexity involved during the compression process.

*Keyewords :* *N-Square, Non-Binary Huffman, Prefix codes.*

*GJCST Classification :* *1.4.2*

AN N-SQUARE APPROACH FOR REDUCED COMPLEXITY NON-BINARY ENCODING

*Strictly as per the compliance and regulations of:*

# An N-Square Approach for Reduced Complexity Non-Binary Encoding

G.Srinivas[α], P.V.G.D.Prasad Reddy[Ω], K.Ramya[β]

*Abstract -* There is always a need for the compression of data to facilitate its easy transmission and storage. Several lossy and lossless techniques have been developed in the past few decades. Lossless techniques allow compression without any loss of information. In this paper, we propose a new algorithm for lossless compression. Our experimental results show that the proposed algorithm performs compression in lesser iterations than the existing Non-Binary Huffman coding without affecting the average number of digits required to represent the symbols, thereby reducing the complexity involved during the compression process.

*Keywords : N-Square, Non-Binary Huffman, Prefix codes.*

## I. Introduction

The basic idea of compression by removal of coding redundancy is to represent the symbols with as less number of bits as possible, so that the average number of bits required to represent the data is reduced. This can be optimized by assigning shorter codes to more frequent symbols and longer codes to less frequent symbols i.e. if A(i) represents $i^{th}$ symbol of the symbol array A, assuming there are a total of n symbols, and let P[A(i)] represent probability of occurrence of the $i^{th}$ symbol,

$$\text{if } P[A(i)] <= P(A(j)] \qquad (1)$$

$$\text{then } L[A(i)] >= L(A(j)] \qquad (2)$$

Where L(A(i)] represents the length of code assigned to the $i^{th}$ symbol. Once the codes are assigned and the symbols are transmitted, the string of code words at the receiving end should be uniquely decodable. To allow unique decodability, we assign prefix codes. Prefix codes are those codes where no code assigned to a symbol is prefix to the code assigned to some other symbol [2],[4].

## II. Existing Method

The optimum coding of message symbols using the Non-Binary Huffman technique is similar to the binary Huffman Coding technique. Every iteration, the least probable D symbols are represented in the next iteration with a new imaginary symbol whose probability equals the sum of probabilities of the constituting symbols. Therefore the number of symbols reduces by D-1 every iteration. To assign 0,1,….,D-1 codes to the symbols remaining in the last iteration, we need to have D symbols left. To allow this,

J symbols are added up in the $1^{st}$ iteration [4]. Where J is given by the formula:

$$J = 2 + [(N-2) \% (D-1)] \quad \text{---------} \ (3)$$

Where N is the total number of symbols in the sample. There after D least probable symbols are added up every iteration till only D symbols are left.

### A. Illustration with an Example

Consider the 18-symbol sample as shown in Table 1. The iterations and the final codes assigned are shown in Table 1(Codes are shown in the brackets). (Assuming the value of D is 5). Here the number of symbols added up in the $1^{st}$ iteration is given by

$$J = 2+[(18-2)\%(5-1)]=2$$

| Original symbol probabilities | Intermediate Iterations | | | Final Iteration- D symbols left |
|---|---|---|---|---|
| 0.1 (3) | 0.1 (3) | 0.18 (2) | 0.22 (1) | 0.4 (0) |
| 0.1 (4) | 0.1 (4) | 0.1 (3) | 0.18 (2) | 0.22 (1) |
| 0.1 (00) | 0.1 (00) | 0.1 (4) | 0.1 (3) | 0.18 (2) |
| 0.1 (01) | 0.1 (01) | 0.1 (00) | 0.1 (4) | 0.1 (3) |
| 0.1 (02) | 0.1 (02) | 0.1 (01) | 0.1 (00) | 0.1 (4) |
| 0.05 (03) | 0.05 (03) | 0.1 (02) | 0.1 (01) | |
| 0.05 (04) | 0.05 (04) | 0.05 (03) | 0.1 (02) | |
| 0.05 (10) | 0.05 (10) | 0.05 (04) | 0.05 (03) | |
| 0.05 (11) | 0.05 (11) | 0.05 (10) | 0.05 (04) | |
| 0.04 (12) | 0.04 (12) | 0.05 (11) | | |
| 0.04 (13) | 0.04 (13) | 0.04 (12) | | |
| 0.04 (14) | 0.04 (14) | 0.04 (13) | | |
| 0.04 (20) | 0.04 (20) | 0.04 (14) | | |
| 0.04 (21) | 0.04 (21) | | | |
| 0.03 (23) | 0.04 (22) | | | |
| 0.03 (24) | 0.03 (23) | | | |
| 0.02 (220) | 0.03 (24) | | | |
| 0.02 (221) | | | | |
| Average Length= | | | | 1.84 |

*Table 1 : Codes assigned using existing Non-Binary Huffman Technique*

## III. Proposed Method

The erratic nature of Non-Binary coding stems from the fact that no pattern has been proposed to chose the value of D. Different D values chosen give different results. One can take the best possible D value to reduce the number of iterations as well as the

*Author : Email : srinivas.gitam@gmail.com, prof.prasadreddy@gmail.com, ramyaa.krothapalli@gmail.com*

number of symbols in each iteration. We propose a pattern to assign a value to D.

*a) N - Square Approach to choose D value*

The number of symbols N determines the D value[10]. D is given by

- Positive square root of N , if N is a perfect square.
- Positive square root of the smallest perfect square greater than N, if N is not a perfect square.

Some N values and their corresponding D values are shown in Table 2.

| N | D |
|---|---|
| 6 | 3 |
| 12 | 4 |
| 23 | 5 |
| 36 | 6 |
| 40 | 7 |
| 56 | 8 |

*Table 2 : 'N' and corresponding 'D' values*

*b) Number of symbols to be added up in the 1st iteration.*

The number of symbols to be added up in the 1st iteration is given by the formula

*P= (D-1)+N% (D-1) ------------(4)*

where N is the total number of symbols in the sample and D is obtained from step A.

*c) Code generation and assignment to the symbols added up in the 1st iteration.*

In cases where the number of symbols added up in the 1st iteration is less than or equal to the D value, the codes are obtained by concatenating 0,1,2…D-1 to the prefix. (The prefix here is the code assigned to the imaginary symbol probability which represents the symbols added up in the 1st iteration). In cases where the number of symbols added up in the 1st iteration exceed the D value, the prefix is concatenated with 0, 1 , 2 …D-2 ,(D-1)0,(D-1)1…. and assigned to the symbols.

*A. Illustration with an example*

Reconsider the 18-symbol example. Since the number of symbols, N=18, D is assigned 5. The number of symbols to be added up in the 1st iteration is given by

*P=(5-1)+18%(5-1)=6.*

Then on, every iteration , 5 symbols are added up. The probability which represents the 6 symbols that are added in the 1st iteration is assigned a code during the process, which is used as prefix to all the 6 symbols. Since the D value is 5, we can concatenate only 5 different symbols to the prefix obtained. In order to generate 6 uniquely decodable codes for the 6 symbols, we expand the obtained 5th code by concatenating 0, 1 to it. In the example prefix 2 is obtained for the 6 symbols. There 20,21,22,23,240,241 are the final 6 codes assigned to the symbols. It is evident that the number of iterations are reduced in the proposed method without effecting the average length of the code. In cases where a large number of symbols is involved, complexity is reduced by a greater degree since the number of symbols reduced per iteration is maximized.

| Original Symbol Probabilities | Intermediate Iterations | | Final Iteration |
|---|---|---|---|
| 0.1 (3) | 0.17 (2) | 0.22 (1) | 0.4 (0) |
| 0.1 (4) | 0.1 (3) | 0.17 (2) | 0.22 (!) |
| 0.1 (00) | 0.1 (4) | 0.1 (3) | 0.17 (2) |
| 0.1 (01) | 0.1 (00) | 0.1 (4) | 0.1 (3) |
| 0.1 (02) | 0.1 (01) | 0.1 (00) | 0.1 (4) |
| 0.05 (03) | 0.1 (02) | 0.1 (01) | |
| 0.05 (04) | 0.05 (03) | 0.1 (02) | |
| 0.05 (10) | 0.05 (04) | 0.05 (03) | |
| 0.05 (11) | 0.05 (10) | 0.05 (04) | |
| 0.04 (12) | 0.05 (11) | | |
| 0.04 (13) | 0.04 (12) | | |
| 0.04 (!4) | 0.04 (13) | | |
| 0.04 (20) | 0.04 (14) | | |
| 0.04 (21) | | | |
| 0.03 (22) | | | |
| 0.03 (23) | | | |
| 0.02 (240) | | | |
| 0.02 (241) | | | |
| Average Length = | | | 1.84 |

*Table 3 :* Codes assigned using the Proposed Method

## IV. DECODING

The codes pertaining to the symbols are sent across from the sender to the receiving end in the form of strings of digits ranging from 0 to D-1. Since the codes assigned are prefix codes[2], the string can be uniquely decodable. A single pass over the string can allow us to determine the unique symbols which correspond to the codes.

## V. CONCLUSION

Experiments show that the proposed method allows assigning codes to the symbols using lesser number of iterations than those employed in the original Non-Binary Huffman Coding. In case of samples where the number of symbols are too many, the number of iterations are considerably reduced thereby reducing the complexity involved.

## REFERENCES RÉFÉRENCES REFERENCIAS

1. S.Jayaram, S.Esakkirajan, T.VeeraKumar , "Digital Image Processing", Tata Mc grawHill Publications.

2. Rafel c. Gonzalez, Richard E.Woods "Digital Image Processing", Second Edition, 6Th Impression 2009, ISBN 978-81-7758-168-3.
3. S.Annadurai, R.ShanmughaLakshmi, "Fundamental of Digital Image processing", 3rd Impression ,ISBN 978-7758-479-0.
4. David A Huffman,Associate,IRE, "A Method for the Construction of Minimum-Redundancy Codes",1952.
5. Cebrail Taskin, Serdar Kürşat Sarikoz, "An Overview of Image Compression Approaches", ICDT,2008.
6. Lu G.,"Advances in digital image compression techniques",1993.
7. [7] Bei Chen,Hong-Cai Zhang,Wen-Lun CAO, Jian-Hu Feng, "Huffman Coding Method based on Number Character", ICMLC 19-22 Aug-2007 IEEE.
8. Reza Hashemian.,"Direct Huffman Coding and Decoding using the table of Code-Lengths",ITCC'03 IEEE.
9. Jia-yu Lin, Ying Liu and ke-Chu Yi, "Balance of 0,1 Bits for Huffman and Reversible Variable-Length Coding", IEEE Transactions on Communications, VOL.52, NO.3, March 2004.
10. G.Srinivas et. al., "N-Square Approach for the Erection of Redundancy Codes", International Journal of Engineering Science and Technology. Vol. 2(3), 2010, 457-461.

This page is intentionally left blank